DOES SYMPLECTIC GEOMETRY YIELD USEFUL NUMERICAL TRICKS?

by

RUA MURRAY
Department of Mathematics and Statistics,
University of Canterbury, Christchurch, New Zealand

No. 110 August, 1994

Abstract

The symplectic structure implicit in systems of Hamilton's equations is of great theoretical, and increasingly numerical, significance. A development of this structure is presented, and ways of utilizing those structural properties in numerical algorithms are discussed. The resulting class of so-called "symplectic integrators" seem to produce good dynamical results. Ways of extending these results to systems of equations with other structural properties are considered.

Chapter 1

Introduction

1.1 Motivation; or why we might look to symplectic geometry for inspiration!

It seems to be a reasonable assumption that a numerical technique for *solving* a system of ordinary differential equations (ODEs) should produce, at least qualitatively, the correct behaviour for that system. In general, such requirements are notoriously difficult to enforce. If, however, the system of ODEs possesses some strong structural[1] properties, then requiring a numerical method to respect these properties might assist in the attainment of reasonably accurate qualitative results.

So called *Hamiltonian* systems possess a lot of structure. This structure is fundamentally connected with the symplectic geometry of the cotangent bundle of the underlying configuration manifold, and can be traced back to the vaguely obscure statement "the τ–flow is symplectic"; what this means (and why I have stated it this way) will be revisited later. For the moment it suffices to consider two tangible consequences:

- preservation of the Hamiltonian (for autonomous systems);
- preservation of phase–space volume.

There are many ways of generating numerical methods [4] that preserve such conditions on an ad hoc basis. The argument for *why* this might be a good idea goes something like this: First, go out and do some analysis on your problem to find some necessary conditions that the solution must satisfy. Convince yourself that making your algorithm satisfy these conditions is a better idea than letting it ignore them. Now, make your algorithm satisfy these conditions. (Conclude with your favourite cynical remark!)

1.2 A 1–F example

Consider the ODE (for a simple conservative pendulum):

$$\begin{aligned}\dot{p} &= -\sin q\\ \dot{q} &= p.\end{aligned}$$

This has a first integral $H = \frac{1}{2}p^2 + 1 - \cos q$. A numerical integrator that gives a "good" solution should preserve this first integral. Hence, examining how H behaves under iterations of the numerical method will give a simple falsification criterion for accuracy.

Following is a picture of a situation in which performance is poor.

[1]The term *structural* is left intentionally vague at this stage!

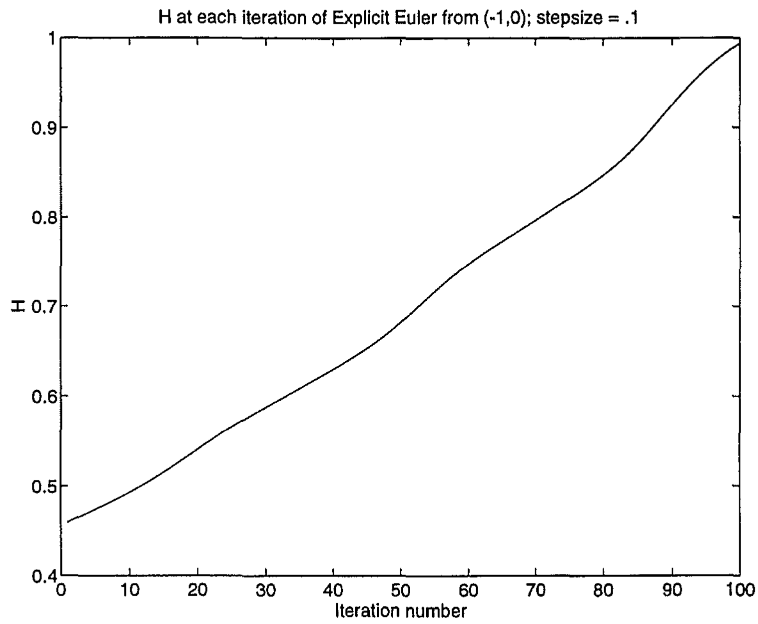

Figure 1 : Euler's method[2] goes awry.[3]

In this case, because the energy surfaces are orbits of the flow, a method which combined a step of Euler's method with a projection onto the energy surface $H = const = H(x_0)$ would retain the correct dynamics; but that is clearly an ad hoc adjustment. It is unlikely that a similar adjustment would achieve anything for an n–F system ($n > 1$). An n–F system is one whose configuration manifold is n– dimensional; since a simple pendulum has the circle S^1 as its configuration manifold, it is a 1–F system.

1.3 Nomenclature

The first step to understanding what is meant by "the τ–flow is symplectic", is to explain what the τ–flow is.

Definition 1.1 *An ODE*

$$\dot{x} = X(x)\,,\, x(0) = x_0$$

has (under appropriate conditions) a solution $x(t) = \phi(t, x_0) = \phi_t(x_0)$. In the case where each ϕ_t is a diffeomorphism and for all s, t in some neighbourhood of 0 $\phi_s \circ \phi_t = \phi_{s+t}$, the set $\{\phi_t\}$ is a one parameter group of diffeomorphisms, and ϕ is the flow *generated by the ODE. The τ–flow is simply the diffeomorphism ϕ_τ.*

The ideal situation in numerical dynamics is that the τ–flow could somehow be shown to stay δ–close to some known map. This map could then be iterated to give the correct qualitative dynamics for the ODE.

Unfortunately, it seems that this situation is a bit much to hope for. Nevertheless, it seems reasonable that approximating the τ–flow by iterations of some known map is the appropriate

[2]If an ODE is given as $\dot{x} = X(x)$ then the *explicit Euler* iteration is just $x_{n+1} = x_n + hX(x_n)$, where h is the fixed *step-size*.

[3]Even if more accurate methods which control local truncation error are employed, the same effect is eventually observed after a large number of time steps.

task for a numerical integrator: comparing the dynamics of a flow with a map of the same dimension does not make very good sense!

Chapter 2 contains a development of the symplectic (differential) geometry necessary to make a global (coordinate free) definition of a system of Hamilton's equations. The symplectic 2–form is defined, and it's invariance under the action of the flow generated by a system of Hamilton's equations is proved. Chapter 3 describes how systems of Hamilton's equations can be obtained via the Legendré transform from certain Lagrangian formulations. This also explains why the natural phase space of a Hamiltonian system is the cotangent bundle (T^*M) of a configuration manifold M, with a consequent symplectic structure. Chapter 4 discusses the practical use of symplectic invariants in the construction of algorithms for the numerical solution of systems of Hamilton's equations. The possibility for extending this technique to systems of equations with other geometric invariants is considered.

Chapter 2

Symplectic geometry and Hamilton's eqns

2.1 Hamilton's equations on $\mathbb{R}^{2n}$

For the moment, consider a system of Hamiltonian (or canonical) equations on an even dimensional Euclidean space: $\mathbb{R}^{2n}$. Let $H : \mathbb{R}^{2n} \to \mathbb{R}$, and let $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n$ so that $\mathbb{R}^{2n}$ has coordinate vector $x = (\mathbf{p}; \mathbf{q})$. Then Hamilton's equations are

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \tag{2.1}$$

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \tag{2.2}$$

where $\frac{\partial H}{\partial \mathbf{p}}$ is the vector with components $(\frac{\partial H}{\partial p^i})_{i=1}^n$ and similarly for $\frac{\partial H}{\partial \mathbf{q}}$. The function H is called the *Hamiltonian*. Note that if, $x = (\mathbf{p}; \mathbf{q})$, then (2.1), (2.2) may be written as $\dot{x} = v_H(x)$ where

$$v_H = \left(-\frac{\partial H}{\partial \mathbf{q}}, \frac{\partial H}{\partial \mathbf{p}} \right) = J\ \nabla_x H, \tag{2.3}$$

J is the $2n \times 2n$ matrix $\left(\begin{smallmatrix} 0 & -I_n \\ I_n & 0 \end{smallmatrix}\right)$[1] and ∇_x is the usual gradient operator on $\mathbb{R}^{2n}$.

Now, with Hamilton's equations in this coordinate form, it is possible to verify the two invariance properties mentioned in the previous section.

Let $x(t) : \mathbb{R} \to \mathbb{R}^{2n}$ be an *integral curve* of a system of canonical equations. Then preservation of H merely means that $H \circ x : \mathbb{R} \to \mathbb{R}$ must be constant with respect to t. This occurs precisely when the system is autonomous since

$$\frac{d(H \circ x)}{dt} = \sum_{i=1}^n \left(\frac{\partial H}{\partial p^i}\dot{p}^i + \frac{\partial H}{\partial q^i}\dot{q}^i \right) + \frac{\partial H}{\partial t} = \sum_{i=1}^n \left(\frac{\partial H}{\partial p^i}\left(-\frac{\partial H}{\partial q^i} \right) + \frac{\partial H}{\partial q^i}\frac{\partial H}{\partial p^i} \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t},$$

which is identically equal to zero whenever H is independent of t.

For volume preservation, observe from (2.3) that $\nabla_x \cdot v_H = \nabla_x \cdot J\ \nabla_x H$. But since the matrix J is antisymmetric, this expression must be zero. Thus the Hamiltonian vector field is divergence free, so the associated flow must preserve volume.

This coordinate–wise formulation of Hamilton's equations is essentially the most rudimentary. The more general, geometric and inspiration yielding approach is in terms of symplectic structures on manifolds. We now pass to such structures, and eventually (theorem 2.9) will show that the above coordinate–wise formulation is reproduced.

[1] I_n is the $n \times n$ identity matrix.

2.2 Symplectic manifolds (the symplectic 2–form)

In this section, only even dimensional manifolds will be considered; symplectic structure can only be defined on even dimensional manifolds. This means restricting attention to *autonomous* systems of canonical equations which must always occur on an even dimensional manifold (see Chapter 3). The general (nonautonomous) canonical equations then require a manifold of odd dimension, on which something analogous to a symplectic structure can be defined ; see Arnold [1, §44].

Let $M = M^{2n}$ be a $2n$ dimensional manifold. A *symplectic* manifold (M^{2n}, ω^2) is an even dimensional manifold, supporting a symplectic structure: the symplectic 2–form ω^2. For the basic facts of differential geometry which are required henceforth, an introductory text such as Spivak [13], would be an excellent reference.

Definition 2.1 *A* symplectic structure *on the manifold M^{2n} is a closed, non-degenerate 2–form ω^2. That is, for u, v vector fields on M^{2n},*

$$d\omega^2 = 0 \quad , \forall u \neq 0 \; \exists v : \omega^2(u,v) \neq 0.$$

$$\omega^2(u,v) = -\omega^2(v,u),$$

Example 2.2 ($\omega^2 = d\mathbf{p} \wedge d\mathbf{q}$) *If M is $\mathbb{R}^{2n}$, then the 2-form $d\mathbf{p} \wedge d\mathbf{q} = \sum_{j=1}^{n} dp^j \wedge dq^j$ provides $\mathbb{R}^{2n}$ with a symplectic structure. The dp^j, dq^j are basis elements for the dual space to $\mathbb{R}^n$.*

The anti–symmetry is obvious from the properties of the wedge product $\wedge$, and the fact that dp^j and dq^j are coordinate differentials implies closure. For the non–degeneracy condition, we make use of the identification of the tangent space at a point of $\mathbb{R}^{2n}$ with $\mathbb{R}^{2n}$, and use the following lemma.

Lemma 2.3 (Algebraic representation of ω^2 on $\mathbb{R}^{2n}$) *Let $x \in M^{2n}$ be at the origin of a coordinate patch $(\mathbf{p}; \mathbf{q})$, and let ω^2 be given as $d\mathbf{p} \wedge d\mathbf{q}$ on $T_x M^{2n} = \mathbb{R}^{2n}$. Then for $\xi, \eta \in T_x M^{2n}$,*

$$\omega^2(\xi, \eta) = \eta^T J \xi \,, \tag{2.4}$$

where $J = \left(\begin{smallmatrix} 0 & -I_n \\ I_n & 0 \end{smallmatrix}\right)$.

Proof: Note that $d\mathbf{p} \wedge d\mathbf{q} = \sum_{j=1}^{n} dp^j \wedge dq^j = \sum_{j=1}^{n} dp^j \otimes dq^j - dq^j \otimes dp^j$. Hence, if $\xi = (\xi^{p_1}, \ldots, \xi^{p_n}, \xi^{q_1}, \ldots, \xi^{q_n})$ and $\eta = (\eta^{p_1}, \ldots, \eta^{p_n}, \eta^{q_1}, \ldots, \eta^{q_n})$,

$$\begin{aligned} \omega^2(\xi, \eta) &= \sum_{j=1}^{n} (dp^j \otimes dq^j - dq^j \otimes dp^j)(\xi, \eta) = \sum_{j=1}^{n} (\xi^{p_j} \eta^{q_j} - \eta^{p_j} \xi^{q_j}) \\ &= (\xi^{p_1}, \ldots, \xi^{p_n}, \xi^{q_1}, \ldots, \xi^{q_n}) \cdot (\eta^{q_1}, \ldots, \eta^{q_n}, -\eta^{p_1}, \ldots, -\eta^{p_n}) \\ &= -\xi^T J \eta = \eta^T J \xi, \end{aligned}$$

where $\cdot$ denotes the normal Euclidean scalar product.■

Now, given $\xi \in \mathbb{R}^{2n}$, putting $\eta = J\xi$ and using (2.4) gives

$$\omega^2(\xi, \eta) = \xi^T J J \xi = \xi^T(-I_{2n})\xi = -\xi \cdot \xi \neq 0$$

whenever $\xi \neq 0$, thus proving the non–degeneracy condition.

Example 2.4 *Let M be an n–manifold with a local coordinate system $\{\mathbf{q}\}$. Let T^*M, the cotangent bundle of M, be coordinatized by $\{(\mathbf{p}; \mathbf{q})\}$ (the $\{\mathbf{p}\}$ are the fibre coordinates). With this system of coordinates, T^*M is locally diffeomorphic to $\mathbb{R}^{2n}$, and the construction of the previous example gives a natural symplectic structure for T^*M.*

In general, the 2–form supported on a $2n$ dimensional symplectic manifold need not have the above form. However, a theorem of *Darboux* (see Arnold [1, §43B]) shows that for any closed non–degenerate 2–form on a neighbourhood of $\mathbb{R}^{2n}$, it is possible to construct a coordinate system $(\mathbf{p};\mathbf{q})$ in which the form is represented as $\omega^2 = d\mathbf{p}\wedge d\mathbf{q}$.

Hence, we can truly regard $d\mathbf{p}\wedge d\mathbf{q}$ as the canonical form of the canonical form!

Note: Standard formulations of the Riemannian geometry of a manifold hinge on the existence of a positive definite, symmetric, bilinear $\binom{0}{2}$–tensor, the *metric.* A manifold equipped with such a metric is called a *Riemannian manifold*, and at each point of the manifold, the action of the metric on each tangent space is similar to that in (2.4) (with the matrix J replaced by some non–singular, symmetric positive definite matrix). The metric tensor can be used in the obvious way to define a scalar product.

For example, equipping $\mathbb{R}^n$ with the standard inner product $\langle\xi,\eta\rangle = \xi\cdot\eta = \xi^T I_n \eta$ yields a Riemannian manifold. The basic geometric properties of $\mathbb{R}^n$ can thus be deduced, for example, to each vector $\eta\in\mathbb{R}^n$ there corresponds an $(n-1)$ dimensional orthogonal complement $(ker\langle\eta,\cdot\rangle)$.

The so–called *symplectic geometry* (of an even dimensional manifold) proceeds by defining a skew–scalar product, $[\xi,\eta] = \xi^T A\eta$ on each tangent space, where each A is a non–singular skew symmetric matrix. The symplectic 2–form ω^2 is one such source a skew–scalar product: $[\xi,\eta] = \omega^2(\xi,\eta)$ (cf. (2.4)) and the properties of symplectic geometry can be deduced. For example, the skew complement of a vector $\xi\in\mathbb{R}^{2n}$ can be defined $(ker[\xi,\cdot])$.

2.3 Building Hamiltonian vector fields ($v_H = IdH$)

An important fact in Riemannian geometry is that the metric tensor can be used to put vector fields in a bijective correspondence with 1–forms. For example, given a vector $\eta\in T_xM^n$ there exists (by the Reisz representation theorem) a unique linear functional f_η such that $f_\eta(\xi) = \langle\xi,\eta\rangle$ for all $\xi\in T_xM^n$. Hence, given a vector field v on M^n, a 1–form ω_v^1 can be defined by putting $\omega_v^1|_x = f_\eta$ where $\eta = v|_{T_xM^n}$ for each $x\in M^n$.

We can use the symplectic 2–form in precisely the same way on a symplectic manifold. This is the approach of Arnold [1, §37C].

Definition 2.5 *Let v be a vector field on a symplectic manifold (M^{2n},ω^2). To each $x\in M^{2n}$ let $\eta = v|_{T_xM^{2n}}$ and let f_η be the unique linear functional which acts on T_xM^{2n} according to*

$$f_\eta(\xi) = \omega^2(\xi,\eta) \quad \forall\, \xi\in T_xM^{2n}.$$

Then $f_\eta\in T_x^*M^{2n}$ for each $x\in M^{2n}$ and let $\omega_v^1\in T^*M^{2n}$ be the 1-form whose restriction to each cotangent space agrees with each f_η.

Theorem 2.6 *Let $\mathbb{R}^{2n} = \{(\mathbf{p};\mathbf{q})\}$ have the standard Euclidean structure. Then, 1-forms can be identified as "row vectors" and vectors as the usual "column vectors". The correspondence $\eta\mapsto\omega_\eta^1$ can be represented as*

$$\tilde{\omega} = \eta^T J,$$

where $\tilde{\omega}$ is the row vector representation of ω_η^1.

Proof: Note that the Euclidean structure of $\mathbb{R}^{2n}$ implies that

$$\omega_\eta^1(\xi) = \langle\xi,\tilde{\omega}^T\rangle = \tilde{\omega}\xi.$$

However, from the definition above and (2.4)

$$\omega_\eta^1(\xi) = \omega^2(\xi,\eta) = \eta^T J\xi = (\eta^T J)\xi.$$

Hence, $\tilde{\omega}\xi = (\eta^T J)\xi$ for all $\xi \in \mathbb{R}^{2n}$; the result follows. ■

Since the matrix J is nonsingular, it is clear that the above transformation in fact defines an isomorphism between between vector fields and 1–forms on $\mathbb{R}^{2n}$. On a manifold M^{2n} an isomorphism can be constructed by identifying each $T_x M^{2n}$ with $T_x^* M^{2n}$ in the above way for each $x \in M^{2n}$. This demonstrates the following proposition.

Proposition 2.7 *Let ω^1 be a 1-form on a symplectic manifold (M^{2n}, ω^2). Then there exists a unique vector field $I\omega^1$ on M^{2n} such that*

$$\omega^1(\cdot) = \omega^2(\cdot, I\omega^1). \tag{2.5}$$

Now, given a function $H : M^{2n} \to \mathbb{R}$, (the *Hamiltonian*) we can take its exterior derivative $d\cdot$ to obtain the 1–form dH.

Definition 2.8 *Let $H : M^{2n} \to \mathbb{R}$ be a differentiable function. To the 1-form dH, there corresponds a unique vector field IdH such that (2.5) is satisfied. This is the* Hamiltonian vector field *$v_H = IdH$, and (2.5) may be rewritten as*

$$dH(\cdot) = \omega^2(\cdot, v_H). \tag{2.6}$$

Theorem 2.9 *Let $H : M^{2n} \to \mathbb{R}$ be a Hamiltonian. Then the vector field $v_H = IdH$ can be given the usual representation:*

$$v_H = \left(-\frac{\partial H}{\partial \mathbf{q}}; \frac{\partial H}{\partial \mathbf{p}} \right)$$

in a patch of canonical coordinates $(\mathbf{p}; \mathbf{q})$.

Proof: Darboux's theorem [1, §43B] guarantees the existence of a system $(\mathbf{p}; \mathbf{q})$ of local coordinates in which ω^2 has the representation $\omega^2 = d\mathbf{p} \wedge d\mathbf{q}$. In this patch, the manifold can be regarded locally as a subset of $\mathbb{R}^{2n}$, so the 1–form dH can be given the row vector representation $(\frac{\partial H}{\partial \mathbf{p}}, \frac{\partial H}{\partial \mathbf{q}})$. Equation (2.6) and theorem 2.6 now imply that

$$\left(\frac{\partial H}{\partial \mathbf{p}}, \frac{\partial H}{\partial \mathbf{q}} \right) = dH(\cdot) = \omega^2(\cdot, v_H) = v_H^T J,$$

or equivalently

$$v_H = J\left(\frac{\partial H}{\partial \mathbf{p}}; \frac{\partial H}{\partial \mathbf{q}} \right) = \left(-\frac{\partial H}{\partial \mathbf{q}}; \frac{\partial H}{\partial \mathbf{p}} \right). \blacksquare$$

Having constructed the Hamiltonian vector field for a Hamiltonian H, it is now essentially a matter of checking definitions to verify that H is in fact invariant under the action of the associated flow.

Recall that if X is a smooth vector field on a manifold M, and $F : M \to \mathbb{R}$ is a smooth function, then if ϕ_t is the flow generated by X,

$$dF(X) = \frac{d}{dt}(F \circ \phi_t)|_{t=0}. \tag{2.7}$$

Theorem 2.10 (Preservation of H) *Let H be a Hamiltonian with associated vector field v_H. Then H is constant under action of the flow generated by v_H.*

Proof: Let ϕ_t be the flow generated by v_H. Then from (2.7) and definition 2.8

$$\frac{d}{dt}(H \circ \phi_t) = dH(v_H) = \omega^2(v_H, v_H).$$

But the rhs of this expression is zero by the skew-symmetry of ω^2, so H is constant on each flow line. ■

It is now clear that conservation of the Hamiltonian is an elementary consequence of the symplectic structure of the underlying manifold; and hence that the symplectic 2-form (the *structural* part of the manifold) is of primary importance. The investigation now turns to the invariance properties of ω^2.

2.4 Preservation of the symplectic 2–form

To begin, it is necessary to establish some notation. Recall that the phase flow associated with a Hamiltonian vector field v_H is denoted by ϕ_t. Preservation of the symplectic 2-form can now be given a precise meaning.

Definition 2.11 (Symplectic maps) *Let (M^{2n}, ω^2) be a sympletic manifold. Then a differentiable map $f : M^{2n} \to M^{2n}$ is* symplectic *iff it preserves the symplectic structure, ie.*

$$f^*\omega^2 = \omega^2$$

where $\cdot^*$ is the usual pull-back operator[2].

Thus, a phase flow is symplectic if $\phi_t^*\omega^2 = \omega^2$ for each ϕ_t.

The proof that the τ–flow is symplectic is accomplished via a slightly more theoretical excursion.

Definition 2.12 (Lie derivative) *Let X be a smooth vector field on a differentiable manifold M. Let ϕ_t be the associated flow, and let ω^p be a smooth p–form on M. Define the* Lie derivative *of $L_X\omega^p$ of ω^p by*

$$L_X\omega^p = \left.\frac{d}{dt}\right|_{t=0} (\phi_t)^*\omega^p. \tag{2.8}$$

Note: For a point $x \in M$, and a p-tuple of vectors $(v_1, \cdots, v_p)$ in T_xM, this formula has the following interpretation:

$$\Big(L_X\omega^p\Big)_x(v_1, \cdots, v_p) = \lim_{t\to 0} \frac{(\phi_t)^*\Big(\omega_x^p(v_1, \cdots, v_p)\Big) - \omega_x^p(v_1, \cdots, v_p)}{t}.$$

Proposition 2.13 *Let X be a smooth vector field on a differentiable manifold M. Let ϕ_t be the associated flow, and let ω^p be a smooth p–form on M. Then $L_X\omega^p = 0$ if and only if $(\phi_t)^*\omega^p = \omega^p$ for each t where ϕ_t is defined.*

Proof: It is easily seen from (2.8) that $L_X\omega^p = 0$ if and only if $(\phi_t)^*\omega^p$ is constant on each orbit of $\{\phi_t\}$. In particular, $(\phi_t)^*\omega^p = (\phi_0)^*\omega^p = \omega^p$.■

It is now useful to obtain a formula for the Lie derivative of a p–form.

[2] *Let ω be a p–form acting on TM (for a manifold M), and $f : N \to M$ a differentiable map where (for another manifold N). Then for a point $q \in N$, the p–form $f^*\omega$ has the action on T_qN of $(f^*\omega)_q(v_1, \cdots, v_p) = \omega_{f(p)}(f_*v_1, \cdots, f_*v_p)$ where $\{v_1, \cdots, v_p\} \subset T_qN$ and $f_* : T_qN \to T_{f(q)}M$ is the tangent map of f.*

Definition 2.14 (Contraction) *Let ω be a p-form on a manifold M. Let X be a vector field on M, and define a $(p-1)$-form $i_X\omega$, the* contraction of ω with X, *by*

$$i_X\omega(v_1, \cdots, v_{p-1}) = \omega(X, v_1, \cdots, v_{p-1})$$

for each $(p-1)$-tuple of vector fields $(v_1, \cdots, v_{p-1})$.

Example 2.15 *Recall from equation (2.6) that the Hamiltonian vector field associated with a Hamiltonian H is defined according to $dH(\cdot) = \omega^2(\cdot, v_H) = -\omega^2(v_H, \cdot)$. Thus*

$$dH = -i_{v_H}\omega^2. \tag{2.9}$$

Lemma 2.16 *Let X be a smooth vector field on a manifold M, and let ω be a differential p-form. Then*

$$L_X\omega = i_X(d\omega) + d(i_X\omega). \tag{2.10}$$

Proof: The proof proceeds by considering the integral of the left and right sides of (2.10) over an arbitrary p-chain c.

The idea is to use the flow ϕ_t, $0 < t < \tau$, to "sweep-out" a $(p+1)$–chain Jc, over which $d\omega$ can be integrated. Let D be a rectangle in $\mathbb{R}^p$ with an oriented frame Or, and let $f : D \to M$ so that (D, f, Or) is a cell in c. Define an associated cell in Jc by $([0,\tau] \times D, f', Or')$ where $f'(t, \mathrm{s}) = \phi_t(f(\mathrm{s}))$ and Or' is the orientation that has the unit vector of the t axis as its first component, and the oriented frame Or as its remaining p components. The boundary of this $(k+1)$–chain consists of the chains[3] c, $\phi_\tau c$ and the trace of ∂c under the action of ϕ_t, $(0 \le t \le \tau)$ (which is just $J\partial c$). Hence, with the orientation Or',

$$\partial(Jc) = \phi_\tau c - c - J\partial c. \tag{2.11}$$

It is sufficient to consider a chain with one cell $f : [0,1]^p \to M$, as a general p–chain is just a linear combination of such cells. Then, letting $f' = \phi_t \circ f$, $\eta = \frac{\partial f'}{\partial t}$ and $\xi_j = \frac{\partial f'}{\partial s^j}$, $(1 \le j \le p)$ the cell $([0,\tau] \times [0,1]^p, f'(t, \mathrm{s}), [\eta, \xi_1, \cdots, \xi_p])$ comprises the chain Jc.

Now, since $X\Big(f'(t, \mathrm{s})\Big) = \frac{d}{dt'}\Big|_{t'=t} \phi_{t'}\Big(f(\mathrm{s})\Big) = \eta$, and $\xi_j = \frac{\partial f'}{\partial s^j} = D\phi_t\Big(f(\mathrm{s})\Big)\frac{\partial f}{\partial s^j}$ we have

$$\begin{aligned} d\omega_{f'(t,\mathrm{s})}(\eta, \xi_1, \cdots, \xi_p) &= d\omega_{f'(t,\mathrm{s})}\Big(X, D\phi_t\big(f(\mathrm{s})\big)\frac{\partial f}{\partial s^1}, \cdots, D\phi_t\big(f(\mathrm{s})\big)\frac{\partial f}{\partial s^p}\Big) \\ &= \Big((\phi_t)^*(i_X d\omega)_{f(\mathrm{s})}\Big)\Big(\frac{\partial f}{\partial s^1}, \cdots, \frac{\partial f}{\partial s^p}\Big). \end{aligned} \tag{2.12}$$

The definition of the integral, equation (2.12) and Fubini's theorem now imply that

$$\begin{aligned} \int_{Jc} d\omega &= \int_{[0,1]^p}\int_0^\tau d\omega(\eta, \xi_1, \cdots, \xi_p)\,dt\,ds_1\cdots ds_p \\ &= \int_{[0,1]^p}\int_0^\tau (\phi_t)^*(i_X d\omega)\Big(\frac{\partial f}{\partial s^1}, \cdots, \frac{\partial f}{\partial s^p}\Big)dt\,ds_1\cdots ds_p \\ &= \int_0^\tau \int_{[0,1]^p} (\phi_t)^*(i_X d\omega)\Big(\frac{\partial f}{\partial s^1}, \cdots, \frac{\partial f}{\partial s^p}\Big)\,ds_1\cdots ds_p\,dt \\ &= \int_0^\tau \int_c (\phi_t)^*(i_X d\omega)dt && (2.13) \\ &= \int_0^\tau \Big(\int_{\phi_t c} i_X d\omega\Big)dt. && (2.14) \end{aligned}$$

[3]If c is a chain with cells $(D \subset \mathbb{R}^n, f : D \to M, Or)$, and $g : M \to M$ is local diffeomorphism of M, then gc is the chain with cells of the form $(D, g \circ f, Or)$.

However, Stokes's theorem and equation (2.11) imply that

$$\int_{Jc} d\omega = \int_{\partial Jc} \omega = \int_{\phi_\tau c} \omega - \int_c \omega - \int_{J\partial c} \omega = \int_c (\phi_\tau)^* \omega - \int_c \omega - \int_{J\partial c} \omega. \tag{2.15}$$

Now, an argument similar to that used to obtain (2.13) can be employed to write

$$\int_{J\partial c} \omega = \int_0^\tau \Big(\int_{\partial c} (\phi_t)^*(i_X \omega) \Big) dt. \tag{2.16}$$

But by Stokes's theorem and the commutativity of $\cdot^*$ and $d\cdot$,

$$\int_{\partial c} (\phi_t)^*(i_X\omega) = \int_c d\Big((\phi_t)^*(i_X\omega)\Big) = \int_c (\phi_t)^* \Big(d(i_X\omega)\Big) = \int_{\phi_t c} d(i_X\omega),$$

so that (2.16) becomes

$$\int_{J\partial c} \omega = \int_0^\tau \Big(\int_{\phi_t c} d(i_X\omega) \Big) dt. \tag{2.17}$$

Equations (2.14), (2.15) and (2.17) can now be combined to give

$$\int_0^\tau \Big(\int_{\phi_t c} i_X d\omega + d(i_X\omega) \Big) dt = \int_c \Big((\phi_\tau)^* \omega - \omega \Big). \tag{2.18}$$

Finally, taking $\frac{d}{d\tau}\Big|_{\tau=0}$ of (2.18) yields,

$$\int_c \Big(i_X d\omega + d(i_X\omega) \Big) = \int_c L_X \omega;$$

from which the result follows since the p–chain c is arbitrary. ■

Theorem 2.17 (The τ–flow is symplectic) *A Hamiltonian flow preserves the symplectic structure: for each τ*

$$(\phi_\tau)^* \omega^2 = \omega^2$$

Proof: Equations (2.10) and (2.9) imply that

$$L_{v_H} \omega^2 = i_{v_H} d\omega^2 + d(i_{v_H} \omega^2) = 0 + d(-dH) = 0$$

since ω^2 is a closed form. The result now follows from proposition 2.13. ■

Let M be a differentiable manifold. A differentiable map $g : M \to M$ has a k–form ω as an *integral invariant* if for every k–chain c on M, $\int_{gc} \omega = \int_c \omega$. A useful characterization of integral invariance is given by the following (rather trivial) proposition.

Proposition 2.18 *A k-form ω is an integral invariant of g if and only if $g^*\omega = \omega$.*

Proof: For an arbitrary k–chain c

$$\int_c g^* \omega = \int_{gc} \omega = \int_c \omega$$

if and only if ω is an integral invariant of g. Since c is arbitary, this occurs precisely when $g^*\omega = \omega$. ■

It is now clear that symplectic maps are a special case of maps possessing integral invariants. Theorem 2.17 and proposition 2.18 thus demonstrate that ω^2 is an integral invariant of the Hamiltonian phase flow.

Lemma 2.19 *Let ω be an integral invariant of g. Then $\omega \wedge \omega$ is also an integral invariant of g.*

Proof: Since $g^*(\omega \wedge \omega) = (g^*\omega) \wedge (g^*\omega)$, proposition 2.18 implies the result. ■

Lemma 2.19 gives a useful way to generate integral invariants for a given differentiable map. In fact, iterated use of lemma 2.19 proves the first part of the following corollary.

Corollary 2.20 (Volume preservation) *Let ϕ_t be a Hamiltonian phase flow on a symplectic manifold (M^{2n}, ω^2). Then the n differential forms*

$$\omega^2 \,, (\omega^2)^2 = \omega^2 \wedge \omega^2 \,, \ldots \,, (\omega^2)^n$$

are integral invariants of ϕ_t. In particular, Hamiltonian phase flows preserve volume.

Proof: Consider the $2n$–form $(\omega^2)^n$. If M^{2n} is $\mathbb{R}^{2n}$, then $(\omega^2)^n$ is simply the usual volume element, and hence for an open set U

$$vol\Big(U\Big) = \int_U (\omega^2)^n = \int_{\phi_t(U)} (\omega^2)^n = vol\Big(\phi_t(U)\Big).$$

For a more general manifold, simply choose $(\omega^2)^n$ to be the volume element, a choice which one is free to make since up to scalar multiples, there is essentially only one candidate $2n$–form. ■

This completes the differential geometric proof of the volume preservation property of Hamiltonian phase flows. It should now be clear that the symplectic 2–form plays a fundamental role in Hamiltonian dynamics. Since preservation of the symplectic 2–form is an intrinsic property of the τ–flow, it seems reasonable to expect that a numerical technique for approximating the τ–flow should also be symplectic. It is thus important to know the answer to the following question: what does a symplectic map look like?

2.5 Symplectic maps have symplectic Jacobians

Theorem 2.21 *Let $g : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a symplectic map. Then at each point $x \in \mathbb{R}^{2n}$, the Jacobian matrix $D_x g$ is a symplectic matrix. That is, $(D_x g)^T J (D_x g) = J$, where $J = \left(\begin{smallmatrix} 0 & -I_n \\ I_n & 0 \end{smallmatrix}\right)$.*

Proof: Since g is a symplectic map, $g^*\omega^2 = \omega^2$. Now, let ξ and η be any two vectors in $T_x\mathbb{R}^{2n} = \mathbb{R}^{2n}$. Then, by the above and the definition of the $\cdot^*$ operator,

$$\omega^2(\xi, \eta) = g^*\omega^2(\xi, \eta) = \omega^2(D_x g\, \xi, D_x g\, \eta).$$

However, recalling the algebraic representation of ω^2 given in equation (2.4), the left and right sides of the above may be rewritten to give

$$\eta^T J \xi^T = (D_x g\, \eta)^T J (D_x g\, \xi) = \eta^T \Big((D_x g)^T J (D_x g)\Big) \xi.$$

Since the vectors ξ and η were arbitrary, this proves the result. ■

Chapter 3

A few tit–bits

The previous chapter has focussed on the construction of a system of Hamiltonian equations, given only the Hamiltonian function H and the symplectic structure of the phase space (manifold). The question arises of *why* symplectic manifolds are the natural configuration spaces for Hamiltonian systems, and of *where* the Hamiltonian function comes from. While this is a digression from the investigation of fundamental invariance properties of Hamiltonian flows, it is nevertheless an interesting aside.

3.1 The Legendré transformation

For many physical problems, the so–called Lagrangian formulation is useful. In fact, the systems of Hamiltonian equations under investigation here can be obtained (under appropriate circumstances) from such a formulation. This is accomplished by a useful tool of convex analysis: the Legendré transformation.

Let $h : U \to \mathbb{R}$ be a smooth convex function, (remember that $U \subset \mathbb{R}^n$ is a convex set). Then it can be seen that for each $\mathbf{x} \in U$ the value $h(\mathbf{x})$ is simply $\sup\{g(\mathbf{x}) | g \leq h, g : U \to \mathbb{R} \text{ is affine}\}$. The graph of the affine function for which this supremum is obtained is simply the tangent hyperplane to the graph of h in $\mathbb{R}^{n+1}$ at x. Let U^* denote the set of all such hyperplanes. Let $\mathbf{y}^*(\mathbf{x}) = \mathbf{y}^* \in U^*$ and define a function $h^* : U^* \to \mathbb{R}$ by $h^*(\mathbf{y}^*) = -\mathbf{y}^*(0)$. The function h^* is then the *convex conjugate* of h. For a fuller discussion, and proofs of the appropriate results, see Rockafeller [10, §12].

A function $h : \mathbb{R}^n \to \mathbb{R}$ and its convex conjugate h^* are geometrically related in the following sense: they both describe the graph of the function h in $\mathbb{R}^n \times \mathbb{R}$. In the case of h, the graph is simply the set of all $(n+1)$ tuples $\big(\mathbf{x}; h(\mathbf{x})\big)$. For a vector $\mathbf{p} \in \mathbb{R}^n$, let $\mathbf{y}^*_{\mathbf{p}}$ be a hyperplane with normal vector $\big(\mathbf{p}; -1\big)$ in $\mathbb{R}^{n+1}$. Then $-h^*(\mathbf{y}^*_{\mathbf{p}})$ provides the intercept with the vertical axis that the hyperplane $\mathbf{y}^*_{\mathbf{p}}$ must have in order for it to be a tangent hyperplane to the graph of h. Since h is convex, the graph of h can then be recovered as the envelope of these tangent hyperplanes.

The function h^* constructed here is in fact the Legendré transform of h.

Definition 3.1 (Legendré transform) *Let $h : U \to \mathbb{R}$ be convex. The* Legendré transform *H of h is given by the formula*

$$H(\mathbf{p}) = \max_{\mathbf{x} \in U} \Big(\langle \mathbf{x}, \mathbf{p} \rangle - h(\mathbf{x}) \Big) \tag{3.1}$$

where $\mathbf{p}$ ranges over the subset of $\mathbb{R}^n$ such that the above is well–defined.

Theorem 3.2 *Let $h : U \to \mathbb{R}$ be smooth and convex. Then $H(\mathbf{p}) = h^*(\mathbf{p})$ for all $\mathbf{p}$ "in U^*".*

Proof: Notation: a vector $\mathbf{p}$ is "in U^*" if there exists some hyperplane in the set U^* of tangent hyperplanes which has the vector $(\mathbf{p}; -1)$ as a normal.

Since the function h is convex and differentiable, there exists a unique value of $\mathbf{x}$ for which the maximum in the rhs of (3.1) is obtained, and this occurs precisely when $\nabla_{\mathbf{x}}\Big(\langle \mathbf{x}, \mathbf{p}\rangle - h(\mathbf{x})\Big) = 0$. That is,

$$\mathbf{p} = \nabla_{\mathbf{x}} h(\mathbf{x}).$$

The vector $\mathbf{p}$ thus obtained is such that $\Big(\mathbf{p}; -1\Big)$ is normal to the graph of h in $\mathbb{R}^{n+1}$, and hence is a normal for the tangent hyperplane to h at $\mathbf{x}$. Furthermore, by the construction of U^* it is clear that to each $\mathbf{p}$ there exists an $\mathbf{x} \in U$ for which this correspondence obtains. Hence, the map $U \ni \mathbf{x} \mapsto \mathbf{p} \in U^*$ is surjective. The convexity of h implies that the correspondence is also injective (the graphs of strictly convex functions are strictly supported by their tangent hyperplanes). We thus obtain a bijective correspondence between U^* and U.

Now, from the construction of the function h^*, the hyperplane in U^* associated with a vector $\mathbf{p}$ is $\{y \in \mathbb{R}^{n+1} | \langle y, \Big(\mathbf{p}; -1\Big)\rangle = h^*(\mathbf{p})\}$[1]. Let $\mathbf{x} \in U$ and let $\mathbf{p} = \nabla_{\mathbf{x}} h(\mathbf{x})$. Then

$$H(\mathbf{p}) = \langle \mathbf{x}, \mathbf{p}\rangle - h(\mathbf{x}) = \langle \Big(\mathbf{x}; h(\mathbf{x})\Big), \Big(\mathbf{p}; -1\Big)\rangle = h^*(\mathbf{p})$$

since the point $\Big(\mathbf{x}; h(\mathbf{x})\Big)$ is obviously on the tangent hyperplane to h at $\mathbf{x}$. ∎

Note that the above proof establishes a bijective correspondence between the elements of U^* and a certain set $\{\mathbf{p}\} \subset \mathbb{R}^n$. Since each $\langle \cdot, \mathbf{p}\rangle$ is a linear functional on $\mathbb{R}^n$, U^* can be regarded as a collection of linear functionals on $\mathbb{R}^n$. Hence, the function $H = h^*$ is actually a function on the dual space to $\mathbb{R}^n$.

Example 3.3 *Let $h : \mathbb{R} \to \mathbb{R}$ be a differentiable convex function. Then for each $p \in \mathbb{R}$ define a line by $\lambda_p = \Big\{px - H(p) \Big| x \in \mathbb{R}\Big\}$, where H is the Legendré transform of h. The function h is the envelope defined by the family of lines parameterized by p.*

Example 3.4 *Let $h(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, \mathbf{x}\rangle$. Then $H(\mathbf{p}) = \frac{1}{2}\langle \mathbf{p}, \mathbf{p}\rangle$.*

Proof: Note that $\nabla_{\mathbf{x}}\Big(\langle \mathbf{x}, \mathbf{p}\rangle - h(\mathbf{x})\Big) = \mathbf{p} - \mathbf{x}$, so that $\mathbf{p} = \mathbf{x}$ and $H(p) = \Big(\langle \mathbf{p}, \mathbf{p}\rangle - h(\mathbf{p})\Big) = \frac{1}{2}\langle \mathbf{p}, \mathbf{p}\rangle$. ∎

From this case, it is apparant that the transform of a quadratic form is itself quadratic. However, while h and H have the same formula, and both act on $\mathbb{R}^n$, it is important to distinguish that the domain of H is actually $\mathbb{R}^n = (\mathbb{R}^n)^*$. This is an important generic property of the general Legendré transformation: it transforms functions on a (normed linear) space, to functions on the dual space. Thus, the Legendré transform gives rise to a duality relation. In fact, the Legendré transformation is involutive. The proof of this is an instructive exercise in the calculation of the transform.

Theorem 3.5 *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a smooth convex function, and H its Legendré transform. Then the Legendré transform of H is h.*

[1]This is just the point–normal equation for a hyperpane: ie. those y such that $\langle y, n\rangle = \langle y_0, n\rangle$ where n is a normal vector to the hyperplane, and y_0 is a point on the plane. Recall from the construction of h^* that $-h^*(\mathbf{p})$ is the intercept with the vertical axis of the hyperplane with tangent $(\mathbf{p}; -1)$; put $y_0 = (\mathbf{0}; -h^*(\mathbf{p}))$.

Proof: Denote the gradient operator $\nabla_{\mathbf{x}}$ by $\frac{\partial}{\partial \mathbf{x}}$. Then

$$H(\mathbf{p}) = \Big(\langle \mathbf{x}(\mathbf{p}), \mathbf{p}\rangle - h\big(\mathbf{x}(\mathbf{p})\big)\Big) \tag{3.2}$$

where $\mathbf{x}(\mathbf{p})$ is the unique $\mathbf{x}$ such that

$$0 = \frac{\partial}{\partial \mathbf{x}}\Big(\langle \mathbf{x}, \mathbf{p}\rangle - h(\mathbf{x})\Big) = \mathbf{p} - \frac{\partial}{\partial \mathbf{x}} h(\mathbf{x}). \tag{3.3}$$

Now, the Legendré transform H^* of H is defined by

$$H^*(\mathbf{y}) = \max_{\mathbf{p}\in\mathbb{R}^n} \langle \mathbf{p}, \mathbf{y}\rangle - H(\mathbf{p}).$$

But this means

$$H^*(\mathbf{y}) = H^*\big(\mathbf{p}(\mathbf{y})\big) = \langle \mathbf{p}(\mathbf{y}), \mathbf{y}\rangle - H\big(\mathbf{p}(\mathbf{y})\big) \tag{3.4}$$

where $\mathbf{p}(\mathbf{y})$ is implicitly defined by $\frac{\partial}{\partial \mathbf{p}}\Big(\langle \mathbf{p}, \mathbf{y}\rangle - H(\mathbf{p})\Big) = 0$. Hence, from (3.2) and the chain rule,

$$\begin{aligned} \mathbf{y} = \frac{\partial}{\partial \mathbf{p}} H(\mathbf{p}) &= \frac{\partial}{\partial \mathbf{p}}\Big(\langle \mathbf{x}(\mathbf{p}), \mathbf{p}\rangle - h\big(\mathbf{x}(\mathbf{p})\big)\Big) \\ &= \mathbf{x}(\mathbf{p}) + \frac{\partial \mathbf{x}}{\partial \mathbf{p}}\mathbf{p} - \frac{\partial \mathbf{x}}{\partial \mathbf{p}}\frac{\partial}{\partial \mathbf{x}} h(\mathbf{x}) \\ &= \mathbf{x}(\mathbf{p}) \end{aligned}$$

where the last equality follows from (3.3). It now follows from (3.4) and the bijective correspondence $\mathbf{p} \leftrightarrow \mathbf{x}$ that

$$H^*(\mathbf{y}) = \langle \mathbf{p}, \mathbf{x}\rangle - H(\mathbf{p}) = h(\mathbf{x}).\blacksquare$$

Note: When taking the Legendré transform of some function, it is not necessary to transform with respect to *all* the variables of that function. Suppose $h : \mathbb{R}^2 \to \mathbb{R}$ where $\mathbb{R}^2 = \Big\{(q; x)\Big\}$. Then let $H : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$H\Big((q;p)\Big) = \max_{x\in\mathbb{R}}\Big(\langle x, p\rangle - h\Big((q;x)\Big)\Big).$$

The function $H\Big((q;\cdot)\Big)$ is just the Legendré transform of $h\Big((q;\cdot)\Big)$ for each $q \in \mathbb{R}$.

3.2 Lagrange's equations on tangent bundle

Many physical systems can be described mathematically by use of the so called *Lagrangian formulation*. Let M be a differentiable n-manifold with the local coordinate system $\Big\{\mathbf{q}\Big\}$. Then the tangent bundle TM has the structure of a differentiable $2n$–manifold with the a local coordinate patch $\Big\{(\mathbf{q}; \dot{\mathbf{q}})\Big\}$ ($\dot{\mathbf{q}}$ denotes a tangent vector to M). Now, physical considerations may lead to the definition of a *Lagrangian* function $L : TM \times \mathbb{R} \to \mathbb{R}$, this is written $L = L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The principle of least action then asserts that the curves $\mathbf{q}(t)$ which are extremals of the functional

$$\Phi(\mathbf{q}) = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)dt$$

are *motions in the Lagrangian system with configuration manifold M and Lagrangian function* L[1, §19].

The calculus of variations can now be used to give the following result.

Proposition 3.6 *Let $L : TM \times \mathbb{R} \to \mathbb{R}$ be a Lagrangian function with configuration manifold M^n. Then the motions of the Lagrangian dynamical system are solutions of the Lagrange equations:*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \mathbf{q}}. \tag{3.5}$$

For a less superficial discussion, see Arnold [1] for example.

3.3 Hamilton's equations on cotangent bundle

Assume that some dynamical system can be given a Lagrangian formulation. Furthermore, assume that the Lagrangian L is a smooth convex function on each coordinate patch $(\mathbf{q}, \dot{\mathbf{q}})$ of the $2n$–manifold TM. Now, for each fixed $\mathbf{q} \in M$, the set $\{(\mathbf{q}, \dot{\mathbf{q}})\} = T_{\mathbf{q}}M$, and let each dual space $T_{\mathbf{q}}^*M$ be coordinatized by $\{(\mathbf{q}, \mathbf{p})\}$ for each fixed $\mathbf{q}$. Then, T^*M has local coordinates $\{(\mathbf{q}; \mathbf{p})\}$. Since each $T_{\mathbf{q}}^*M$ is the dual space to each $T_{\mathbf{q}}M$, the Legendré transformation applied to $L(\mathbf{q}, \cdot, t)$ will yield a convex function H on $T^*M \times \mathbb{R}$:

$$H(\mathbf{q}, \mathbf{p}, t) = \max_{\dot{\mathbf{q}} \in T_{\mathbf{q}}M} \Big(\langle \dot{\mathbf{q}}, \mathbf{p}\rangle - L\big(\mathbf{q}, \dot{\mathbf{q}}, t\big)\Big). \tag{3.6}$$

Let this function be a Hamiltonian.

It turns out that the Hamiltonian thus defined has precisely the same associated vector field as is obtained from the Lagrangian formulation.

Theorem 3.7 *Let L be a convex Lagrangian on TM^n. Then the Legendré transform H is a convex function on T^*M^n, and the Lagrangian motions $\mathbf{q}(t)$ which satisfy (3.5) also satisfy*

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}, \; \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}$$

where $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$.

Proof: Recall from (3.6) that

$$H(\mathbf{q}, \mathbf{p}, t) = \langle \dot{\mathbf{q}}(\mathbf{p}), \mathbf{p}\rangle - L\big(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{p}), t\big),$$

where $\dot{\mathbf{q}}(\mathbf{p})$ is implicitly defined by

$$0 = \frac{\partial}{\partial \dot{\mathbf{q}}}\Big(\langle \dot{\mathbf{q}}, \mathbf{p}\rangle - L\big(\mathbf{q}, \dot{\mathbf{q}}, t\big)\Big) = \mathbf{p} - \frac{\partial L}{\partial \dot{\mathbf{q}}}. \tag{3.7}$$

Now, consider the exterior derivative of H. This is defined by

$$dH = \frac{\partial H}{\partial \mathbf{p}} d\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial H}{\partial t} dt. \tag{3.8}$$

However, since H as the Legendré transform of L, dH can be calculated directly:

$$dH = \Big(\frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{p}}\mathbf{p} + \dot{\mathbf{q}}(\mathbf{p}) - \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{p}}\frac{\partial L}{\partial \dot{\mathbf{q}}}\Big) d\mathbf{p} + \Big(-\frac{\partial L}{\partial \mathbf{q}}\Big) d\mathbf{q} - \frac{\partial L}{\partial t} dt \tag{3.9}$$

But comparing equations (3.8) and (3.9), and using equations (3.5) and (3.7), the following identities are obtained:

$$\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}} \;\;, \frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}} = -\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{\mathbf{q}}}\Big) = -\frac{d}{dt}\mathbf{p} = -\dot{\mathbf{p}} \;\;, \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \blacksquare$$

Thus, a dynamical system on a manifold M^n which is formulated in terms of a Lagrangian on TM has a natural Hamiltonian counterpart on the $2n$–dimensional manifold T^*M. This explains why Hamiltonian systems appear on even dimensional manifolds: the phase space is in fact the cotangent bundle of some configuration manifold, and is thus a manifold of dimension twice that of the configuration manifold.

Example 3.8 (Classical mechanics) *Consider a system of particles interacting via some (time invariant) conservative force. Then there exists some potential energy function $V(q)$ (q is simply a coordinate system describing the location of the particles: for m particles on an n dimensional configuration manifold, q has mn components). Also, define a kinetic energy function by $T(\dot{q}) = \frac{1}{2}\langle \dot{q}, \dot{q}\rangle$, for some appropriately chosen inner product. Note that the elements $\dot{q}$ lie in the tangent spaces to M. Put $L(q, \dot{q}) = T(\dot{q}) - V(q)$, so that L is a Lagrangian function on the tangent bundle TM to the configuration manifold. Then, Lagrange's equations (3.5) imply that $\ddot{q} = -\frac{\partial V}{\partial q}$; these are the equations of motion.*

The Legendré transform H of L is

$$H(q,p) = \langle \dot{q}(p), p\rangle - L\Big(q, \dot{q}(p)\Big) = \langle p, p\rangle - \Big(T(p) - V(q)\Big) = T(p) + V(q),$$

since $p = -\frac{\partial L}{\partial \dot{q}} = \frac{\partial}{\partial \dot{q}}\frac{1}{2}\langle \dot{q}, \dot{q}\rangle = \dot{q}$. The Hamiltonian H thus represents total energy. Theorem 2.10 then says that the total energy of the system is conserved.

Example 3.9 (Simple pendulum) *Consider a unit mass suspended by a rod of unit length. Suppose a force acting on the mass has potential energy function $V(q) = 1 - \cos q$ where q is the angle of inclination from the stable equilibrium point (the force $\frac{\partial V}{\partial q}$ represents a constant gravitational force, q is the clockwise angle that the pendulum is from vertical). Hence, the configuration manifold is the circle S^1. Using the above construction, the Hamiltonian $H(q,p) = T(p) + V(q) = \frac{1}{2}\langle p, p\rangle + 1 - \cos q$ represents the total energy of the pendulum, and is a function on T^*S^1. The dynamical equations can now be derived in the standard way.*

Chapter 4

Invariants for algorithms

4.1 Invariants of a flow

Let ϕ_t be the flow associated with X, a smooth vector field on a differentiable manifold M^n. Let ω be a p–form on M. Then propositions 2.13 and 2.18 can be summarized in the following proposition.

Proposition 4.1 *The p–form ω is an integral invariant (invariant) of the flow ϕ_t if and only if $(\phi_t)^*\omega = \omega$ for each t, if and only if $L_X\omega = 0$.*

Corollary 4.2 *Let v_H be a Hamiltonian vector field. Then H, ω^2 and volume are invariants of the Hamiltonian flow.*

Proof: Observe from equations (2.8) and (2.7) that theorem 2.10 can be rewritten as $L_{v_H}H = 0$; since H is a 0–form, the result for H follows from the above proposition. The invariance of ω^2 and *vol* is simply theorem 2.17 and corollary 2.20.■

We have thus obtained a tangible condition for invariance: namely, $L_X \cdot = 0$. The above corollary also offers several invariants of a Hamiltonian flow (more can be obtained from lemma 2.19).

From the point of view of actually constructing implementable algorithms, it is an absolute necessity to express this condition in some form which is algebraic with respect to some coordinate system; rather than with respect to the exterior differential calculus!

4.2 Gear's approach

The conceptual basis of this section is found in a paper of Gear [6]. The expression of these concepts in terms of the Lie derivative is more similar to the approach of Mackay [7].

The observation that H can be regarded as a 0–form has significance from the point of view of verifying integral invariance[1]. Consider an algorithm which approximates the τ–flow of a Hamiltonian system. It is easy to check whether or not this algorithm preserves H: simply evaluate the function H at each point (iterate) generated by the algorithm; this is the technique that was used to generate figure 1.

Measuring the extent to which an algorithm preserves a p–form is not such an easy task: to check invariance of a p–form, one must compare ω to $(\phi_{nh})^*\omega$ (where n is the number of the iterate to be compared, and h the step–size of the algorithm). But to know what $(\phi_t)^*\omega$ is for some t presupposes exact knowledge of the behaviour of ϕ_t! Hence, *preserving* differential invariants is likely to be a difficult task.

[1]That is, it is the only invariant whose invariance can be monitored.

Let ω be a p-form on a manifold M^n. Then each T_x^*M has a basis $\{dx^1, \cdots, dx^n\}$ and ω may be written as

$$\omega = \sum_{1 \le i_1 < \cdots < i_p \le n} \omega_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

where each $\omega_{i_1 \cdots i_p}$ is a smooth real valued function of $x \in M$ in each coordinate patch. This may be rewritten as

$$\omega = \sum_{i_1, \cdots, i_p = 1}^{n} A_{i_1 \cdots i_p} dx^{i_1} \otimes \cdots \otimes dx^{i_p}$$

where the n^p coefficients $A_{i_1 \cdots i_p}$ are recovered from the $\binom{n}{p}$ coefficients $\omega_{i_1 \cdots i_p}$ by the definition of $\wedge$.

Now, to compute the Lie derivative of such a form, it is first necessary to establish a bit more notation. Let $x \in M$ be in some coordinate patch $\left\{(x^1, \cdots, x^n)\right\}$, and let T_xM have a basis $\{\partial_1, \cdots \partial_n\}$. Then a vector $v \in T_xM$ has coordinate representation $v = \sum_{j=1}^n v^j \partial_j$. Recall from standard dual space theory that the dx^i's and ∂_j's satisfy the duality relation

$$dx^i(\partial_j) = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}.$$

It follows from linearity that

$$dx^i(v) = \sum_{j=1}^{n} v^j dx^i(\partial_j) = v^i. \tag{4.1}$$

Lemma 4.3 *Let M be a manifold, and X a smooth vector field on M. Then in each coordinate patch,*

$$L_X dx^i = \sum_{j=1}^{n} \frac{\partial X^i}{\partial x^j} dx^j \tag{4.2}$$

where X has coordinate representation $X = \sum_{j=1}^n X^j \partial_j$.

Proof: We use equation (2.10), the fact that $d(dx^i) = 0$ (dx^i is a coordinate differential) and the standard formula for $d\cdot$ to obtain

$$L_X dx^i = i_X d(dx^i) + d(i_X dx^i) = 0 + d(dx^i(X)) = d(X^i) = \sum_{j=1}^{n} \frac{\partial X^i}{\partial x^j} dx^j. \blacksquare$$

We can now calculate $L_X \omega$.

Theorem 4.4 *The Lie derivative of a p-form is given by*

$$\begin{aligned} L_X \omega &= \sum_{i_1, \cdots, i_p = 1}^{n} (L_X A_{i_1 \cdots i_p}) dx^{i_1} \otimes \cdots \otimes dx^{i_p} \\ &+ \sum_{i_1, \cdots, i_p = 1}^{n} A_{i_1 \cdots i_p} \sum_{j=1}^{n} \frac{\partial X^{i_1}}{\partial x^j} dx^j \otimes \cdots \otimes dx^{i_p} \\ &+ \cdots + \sum_{i_1, \cdots, i_p = 1}^{n} A_{i_1 \cdots i_p} dx^{i_1} \otimes \cdots \otimes \sum_{j=1}^{n} \frac{\partial X^{i_p}}{\partial x^j} dx^j. \end{aligned} \tag{4.3}$$

Proof: Recalling the coordinate–wise interpretation of the definition of the Lie derivative which followed its definition, it can be seen that the usual Leibniz rule applies. This fact, and (4.2) imply the result. $\blacksquare$

Corollary 4.5 *If the p-form ω is such that the functions $A_{i_1\cdots i_p}$ are constant, then $L_X\omega = 0$ if and only if*

$$\sum_{k=1}^{p}\sum_{j=1}^{n} A_{i_1\cdots i_{k-1}ji_{k+1}\cdots i_p}\frac{\partial X^j}{\partial x^{i_k}} = 0$$

for each $\{i_1, \cdots, i_p\}_1^n$.

Proof: If each $A_{i_1\cdots i_p}$ is constant, then $L_X A_{i_1\cdots i_p} = 0$ for each $\{i_1, \cdots, i_p\}_1^n$. The result of theorem 4.4 can then be rewritten as

$$\sum_{i_1,\cdots,i_p=1}^{n}\ \sum_{k=1}^{p}\sum_{j=1}^{n} A_{i_1\cdots i_{k-1}ji_{k+1}\cdots i_p}\frac{\partial X^j}{\partial x^{i_k}} dx^{i_1}\otimes\cdots\otimes dx^{i^p} = 0.$$

The result follows. ∎

Consider now the specific case when ω is a 2–form. Then, in the notation of the above corollary, $p = 2$, and the coefficients $A_{i_1i_2}$ can be regarded as the entries of an $n \times n$ matrix. For each of the n^2 possible choices of (i_1, i_2), the corresponding equations in corollary 4.5 can be interpreted as a way of expressing the (i_1, i_2)th element of the $n \times n$ zero matrix. That is,

Lemma 4.6 *Let $\omega = \sum_{i_1i_2=1}^{n} A_{i_1,1_2}dx^{i_1}\otimes dx^{i_2}$ be a 2-form such that each $L_X A_{i_1i_2} = 0$. Then $L_X\omega = 0$ if and only if*

$$(DX)^T A + A(DX) = 0$$

where DX is the Jacobian matrix of the vector field X, and A is the matrix with entries $A_{i_1i_2}$.

Theorem 4.7 *Let ω^2 be the symplectic 2–form, and let v_H be a Hamiltonian vector field. Then*

$$(Dv_H)^T J + J(Dv_H) = 0$$

where $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$.

Proof: Choose a patch of canonical coordinates $(\mathbf{p}, \mathbf{q})$ so that $\omega^2 = d\mathbf{p} \wedge d\mathbf{q}$. Recall from corollary 4.2 that $L_{v_H}\omega^2 = 0$. Note also that $(x^1, \cdots, x^n, x^{n+1}, \cdots, x^{2n}) = (p^1, \cdots, p^n, q^1, \cdots, q^n)$ so

$$\omega^2 = \sum_{j=1}^{n}(dp^j \otimes dq^j - dq^j \otimes dp^j) = \sum_{j=1}^{n} dx^j \otimes dx^{n+j} - \sum_{j=n+1}^{2n} dx^j \otimes dx^{j-n} = \sum_{i_1,1_2=1}^{2n} A_{i_1i_2}dx^{i_1}\otimes dx^{i_2}.$$

Therefore, the matrix A is simply J, and each $L_{v_H}A_{i_1i_2} = 0$. The result now follows from the above lemma. ∎

4.3 Lie theory

In fact there is nothing new in theorem 4.7. Let v_H be the Hamiltonian vector field with Hamiltonian H on a symplectic manifold (ω^2, M^n). Theorems 2.17 and 2.21 together imply that $\left(D(\phi_t)\right)^T J\left(D(\phi_t)\right) = J$ for each t. This merely reflects the fact that each $D(\phi_t)$ is an element of the Lie group of symplectic matrices: $Sp_{2n}(\mathbb{R}) = \{B \in GL_{2n}(\mathbb{R}) \big| B^T JB = J\}$. Hence, for each $x \in M$, the mapping $t \mapsto \gamma(t) = D(\phi_t)$ defines a curve in $Sp_{2n}(\mathbb{R})$. Since $\phi_0(x) = x$ for each x, $D(\phi_0) = I_{2n}$ so that the tangent space to $Sp_{2n}(\mathbb{R})$ at $D(\phi_0)$ is the Lie algebra $sp_{2n}(\mathbb{R}) = \{B \in GL_{2n}(\mathbb{R}) \big| B^T J + JB = 0\}$. Therefore, the tangent to the curve γ in $Sp_{2n}(\mathbb{R})$ must be an element of $sp_{2n}(\mathbb{R})$. But this tangent is simply

$$\frac{d}{dt}\Big|_{t=0} D(\phi_t) = D\Big(\frac{d}{dt}\Big|_{t=0}(\phi_t)\Big) = D(v_H),$$

so that the result of theorem 4.7 could have been predicted from the earlier result.

This does not mean that corollary 4.5, from which theorem 4.7 was derived contains nothing new. On the contrary, it provides a framework in which to view the general properties of invariants: theorem 4.7 deals only with flows that preserve the symplectic 2–form, whereas corollary 4.5 offers a criterion for the expression of any invariant p–form of any flow.

4.4 Preserving differential invariants

4.4.1 Symplectic algorithms via generating functions

Specializing again to the case of Hamiltonian systems, the question of *how* to preserve the symplectic 2–form becomes pertinent. (For the moment, I shall ignore the problem of preserving the Hamiltonian, for reasons outlined in the section 4.5.)

It is worthwhile to recall what is meant by a numerical integrator. Essentially, for a vector field X on phase space M, a *numerical integrator* with constant step–size τ is a map (preferably diffeomorphic) $g : M \to M$; then[2] $g(x)$ is the numerical approximation to $\phi_\tau(x)$ where $\{\phi_t\}$ is the flow associated with the vector field X. Under such a scheme, given a starting point x_0, $x_n = g^n(x_0)$ is the numerical approximation to $\phi_{n\tau}(x_0)$. The task of constructing a suitable numerical scheme to "integrate" $\dot{x} = X(x)$ then consists in obtaining a function g whose iterates produce an approximation to the orbit of the τ–flow of the flow generated by X. A symplectic integrator is a numerical integrator such that the map g is a symplectic map.

Fortunately, the theory of symplectic maps (or equivalently, canonical transformations) is well developed. See for instance Arnold [1, §47]. There is a bijective correspondence between canonical transformations and so–called generating functions (which are solutions of the Hamilton–Jacobi equation [1, §46]). Perhaps the simplest example of this is drawn from Mackay [7]:

Example 4.8 *Let $S : T^*M \to \mathbb{R}$ be a smooth function $(\mathbf{p}', \mathbf{q}) \mapsto S(\mathbf{p}', \mathbf{q})$, so that for small τ, the relations*

$$\mathbf{p}' = \mathbf{p} - \tau\frac{\partial S}{\partial \mathbf{q}}\big(\mathbf{p}', \mathbf{q}\big), \quad \mathbf{q}' = \mathbf{q} + \tau\frac{\partial S}{\partial \mathbf{p}'}\big(\mathbf{p}', \mathbf{q}\big)$$

implicitly define a map $f_\tau : (\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{p}', \mathbf{q}')$. It can be shown that this map is symplectic.

The most simple-minded application of example 4.8 is simply to let $S = H$, the Hamiltonian of the system. This gives the implicitly defined scheme:

$$(\mathbf{p}_+; \mathbf{q}_+) = (\mathbf{p}; \mathbf{q}) + \tau X\Big((\mathbf{p}_+; \mathbf{q})\Big).$$

Applying this to the example of section 1, the following picture is obtained:

[2] More generally, the map g may actually be chosen at each iteration, but for illustrative purposes there is no utility in considering this generality.

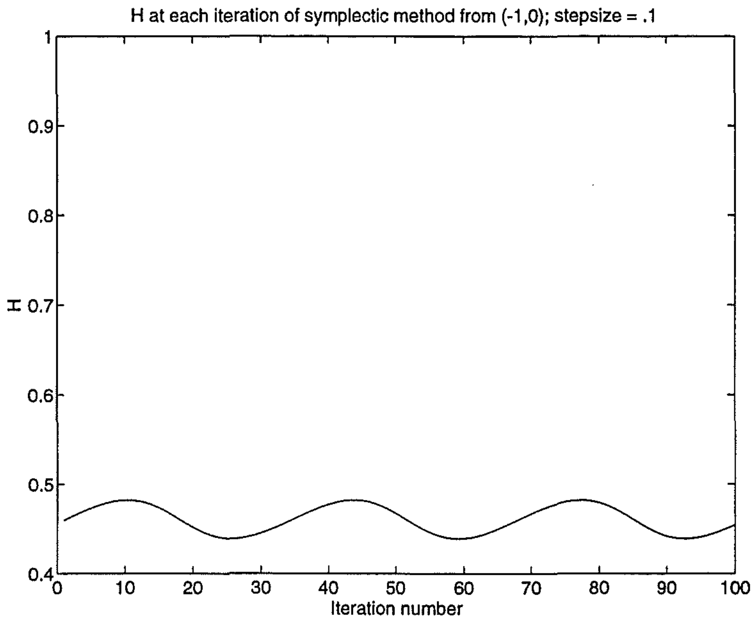

Figure 2 : Observe that this plot of H against iterations is much better behaved than that depicted in figure 1 (a step–size of $0 \cdot 1$ was used here also). While the true dynamics of this flow would produce a picture for which H was constant, the main qualitative feature (periodicity) of the trajectory chosen (the one with initial point $(-1,0)$) is reproduced. Contrast this with the divergence illustrated by figure 1.

It should be clear that in this case at least, a simple symplectic numerical method does give dynamically relevant results. Obviously, far more testing than this is needed, but this simple example is indicative of the kind of result that can be expected. See Mackay [7] for less elementary ways to obtain symplectic integrators from generating functions, and also for some numerical results (and their dynamical implications!). For some more results, see [9].

4.4.2 Symplectic Runge–Kutta

While all symplectic maps do possess generating functions [1], it is possible to have symplectic algorithms which are more general than the basic iteration of a map idea. A broad class of Runge–Kutta algorithms (which consist of a number of internal stages involving the solution of non–linear algebraic equations) also give rise to symplectic transformations of phase–space.

For a discussion of such algorithms, and further references, see Sanz-Serna's paper [11].

4.5 So what

4.5.1 Preserve everything?

The references mentioned above all tend to suggest that using symplectic methods to numerically intergrate systems of Hamilton's equations yield "good" results. It may seem natural to ask whether incorporating other invariance properties (such as conservation of energy and momentum where appropriate) will lead to even better results. The answer to this is yes, but in a useless sense. Suppose f_τ is a map which approximates ϕ_τ, and that f_τ is symplectic and possesses all the symmetries (conserved quantities) of the Hamiltonian phase flow $\{\phi_t\}$. Then Marsden shows [8, p176] that (up to a time reparameterization) the dynamics of the algorithm based on f_τ are identical to the dynamics of the τ–flow. Therefore, such an algorithm will clearly give the correct dynamics, but to construct such an algorithm effectively involves finding the τ–flow, thus pre–empting the need for having an algorithm in the first place!

4.5.2 Shadowing

The ideal situation for the numerical simulation of a phase flow is something like the following. Let the phase flow be given by ϕ, and let $\epsilon > 0$. We would like to be able to find $\tau > 0$ and a corresponding algorithm (regard an algorithm as a method of incrementing $x_n \mapsto x_{n+1}$) such that for each $n > 0$, $||\phi_{n\tau}(x_0) - x_n|| < \epsilon$ for starting points x_0. Since $\phi_{n\tau} = (\phi_\tau)^n$, such a requirement starts to look very like the conclusion of Bowen's shadowing lemma [2, §3].

The shadowing result is proved by methods from ergodic theory. If the eventual aim in numerical intergration is to obtain algorithms for which results similar to the shadowing lemma can be proved, then it is reasonable to expect that some very general method of proof will be required (for example, techniques of ergodic theory). Such proofs will inevitably use structural properties of the τ–flow under consideration. In the case of systems of Hamilton's equations, the fundamental structural property of the τ–flow is its symplectic nature and the consequent (corollary 2.20) volume preservation. Thus, the τ–flow is a volume preserving transformation of phase space, and many results (such as Poincaré's recurrence theorem) apply. A numerically computed orbit which "shadows" the true orbit (the orbit of the τ–flow) must share these properties. This goes some way towards explaining why the numerical use of symplectic (and hence volume preserving) maps is producing such promising results.

4.5.3 Variable step size?

Many successful numerical routines employ variable step sizes to control local truncation error. It might be interesting to know what would happen if variable step size symplectic methods could be designed. A brief paper of Skeel and Gear [12] demonstrates that such interest will inevitably remain unfulfilled. They prove that if the step size is allowed to vary in an algorithm in such a way that retains its symplectic character, then it can only vary slowly (and so will be of little practical use). This should not cause much concern, since it is not at all clear what a variable step size numerical integrator is actually approximating: it is certainly not the τ-flow (since τ would then be able to vary from step to step), and it does not seem to make a great deal of sense from a dynamical point of view to say that the numerically computed (discrete) trajectory approximates an actual (continuous) orbit of the system in any rigorous way.

4.6 Generalizations

Recall that corollary 4.5 provides a collection of algebraic conditions which guarantee the invariance of a p–form under the action of a flow. A similar derivation will provide a characterization for the invariance of an $\binom{n}{m}$–tensor under the action of a flow. More generally still, tensor differential equations can be given an algebraic local coordinate representation.

For example, Crampin [5] discusses a particular $\binom{1}{1}$–tensor S which has fundamental significance for the geometry of the tangent bundle of a differentiable manifold. He proves that for a particular class of vector fields Γ on TM[3], $(L_\Gamma S)^2 = I$ where I denotes the $\binom{1}{1}$ identity tensor[4][5].

This can be converted to a matrix equation which must be satisfied by the Jacobian $D\Gamma = \frac{d}{d\tau}|_{\tau=0}\psi_\tau$ where ψ_t is the one parameter group of diffeomorphisms of TM generated by the vector field Γ. The language of Lie algebras may again be employed to describe that matrix equation as a tangency condition. Then, for any matrix A satisfying that tangency condition,

[3] Each $\Gamma(q, u) \in T_{(q,u)}(TM)$ where (q, u) are the bundle coordinates of TM.

[4] $I(X) = X$ for all vector fields X on TM.

[5] For a $\binom{1}{1}$– tensor T, T^2 denotes the $\binom{1}{1}$–tensor which has the action on vector fields: $((T)^2)(X) = T(T(X))$.

the matrix $B = \exp(A)$ is an element of $GL_{2n}(\mathbb{R})$. Let U be the subset of all $B \in GL_{2n}(\mathbb{R})$ of this form. Let $\mathcal{U} = \{f : TM \to TM \mid Df \in U\}$. Then each $\psi_\tau \in \mathcal{U}$, and using maps $f \in \mathcal{U}$ to approximate ψ_τ could very well result in the same kind of improvements that can be obtained by using symplectic maps of T^*M to approximate the τ–flow of systems of Hamilton's equations.

Crampin also proves that the vector field Γ on TM which results from a system of Lagrange's equations, is of the required form. Thus, pursuing the strategy described in the previous paragraph could very well lead to numerical methods which give good dynamical results for Lagrangian systems. This would be a tremendous advance because only systems with convex Lagrangians can be put into a Hamiltonian form, thus allowing symplectic integrators to be used. Therefore, the possibilities for capitalizing numerically on geometric properties of systems of equations seem to extend far beyond symplectic integrators for Hamiltonian systems.

Bibliography

[1] V I Arnold. *Mathematical methods of classical mechanics.* Springer–Verlag, 2 edition, 1989.

[2] R Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Springer–Verlag, 1970.

[3] D S Broomhead and A Iserles. *The dynamics of numerics and the numerics of dynamics.* Oxford, 1992. Proceedings of IMA conference on dynamics of numerics and numerics of dynamics, Bristol 1990.

[4] J C Butcher. *The numerical analysis of ordinary differential equations.* Wiley, 1987.

[5] M Crampin. Tangent bundle geometry for Lagrangian dynamics. *J. Phys A: Math. Gen.*, 16:3755–3772, 1983.

[6] C W Gear. Invariants and numerical methods for ODEs. *Physica D*, 60:303–310, 1992.

[7] R S Mackay. Some aspects of the dynamics and numerics of Hamiltonian systems. *The dynamics of numerics and the numerics of dynamics [3]*, pages 137–194, 1992.

[8] J E Marsden. *Lectures on mechanics.* Cambridge, 1992.

[9] D Okunbor. Canonical methods for Hamiltonian systems: numerical experiments. *Physica D*, 60:314–322, 1992.

[10] R T Rockafeller. *Convex analysis.* Princeton, 1970.

[11] J M Sanz-Serna. Symplectic Runge–Kutta and related methods: recent results. *Physica D*, 60:293–302, 1992.

[12] R D Skeel and G W Gear. Does variable step–size ruin a symplectic integrator? *Physica D*, 60:311–313, 1992.

[13] M Spivak. *Calculus on manifolds.* W A Benjamin, 1965.